



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/31, 15/63, 15/00, C12P 21/02</b>		<b>A1</b>	(11) International Publication Number: <b>WO 96/33276</b>
			(43) International Publication Date: 24 October 1996 (24.10.96)
(21) International Application Number: <b>PCT/US96/05320</b>		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 22 April 1996 (22.04.96)		<b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data:			
08/426,787      21 April 1995 (21.04.95)      US 08/476,102      7 June 1995 (07.06.95)      US 08/487,429      7 June 1995 (07.06.95)      US			
(71) Applicants: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). JOHNS HOPKINS UNIVERSITY [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US).			
(72) Inventors: FLEISCHMANN, Robert, D.; 470 Tschiffely Square Road, Gaithersburg, MD 20878 (US). ADAMS, Mark, D.; 15205 Dufief Drive, N. Potomac, MD 20878 (US). WHITE, Owen; Apartment #202, 886 Quince Orchard Boulevard, Gaithersburg, MD 20878 (US). SMITH, Hamilton, O.; 8222 Carrbridge Circle, Towson, MD 21204 (US). VENTER, J., Craig; 11915 Glen Mill Road, Potomac, MD 20854 (US).			
(74) Agents: GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, Washington, DC 20005-3934 (US).			
(54) Title: NUCLEOTIDE SEQUENCE OF THE HAEMOPHILUS INFLUENZAE Rd GENOME, FRAGMENTS THEREOF, AND USES THEREOF			
(57) Abstract			
<p>The present invention provides the sequencing of the entire genome of <i>Haemophilus influenzae</i> Rd, SEQ ID NO:1. The present invention further provides the sequence information stored on computer readable media, and computer-based systems and methods which facilitate its use. In addition to the entire genomic sequence, the present invention identifies over 1700 protein encoding fragments of the genome and identifies, by position relative to a unique <i>Not I</i> restriction endonuclease site, any regulatory elements which modulate the expression of the protein encoding fragments of the <i>Haemophilus</i> genome.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

will contain an effective amount of one or more of the agents of the present invention, together with a suitable amount of carrier vehicle.

Additional pharmaceutical methods may be employed to control the duration of action. Control release preparations may be achieved through the use of polymers to complex or absorb one or more of the agents of the present invention. The controlled delivery may be exercised by selecting appropriate macromolecules (for example polyesters, polyamino acids, polyvinyl, pyrrolidone, ethylenevinylacetate, methylcellulose, carboxymethylcellulose, or protamine, sulfate) and the concentration of macromolecules as well as the methods of incorporation in order to control release. Another possible method to control the duration of action by controlled release preparations is to incorporate agents of the present invention into particles of a polymeric material such as polyesters, polyamino acids, hydrogels, poly(lactic acid) or ethylene vinylacetate copolymers. Alternatively, instead of incorporating these agents into polymeric particles, it is possible to entrap these materials in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatine-microcapsules and poly(methylmethacrylate) microcapsules, respectively, or in colloidal drug delivery systems, for example, liposomes, albumin microspheres, microemulsions, nanoparticles, and nanocapsules or in macroemulsions. Such techniques are disclosed in *Remington's Pharmaceutical Sciences* (1980).

The invention further provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the agents of the present invention may be employed in conjunction with other therapeutic compounds.

THIS PAGE BLANK (USPTO)

present invention can be administered concurrently with, prior to, or following the administration of the other agent.

The agents of the present invention are intended to be provided to recipient subjects in an amount sufficient to decrease the rate of growth (as defined above) of the target organism.

The administration of the agent(s) of the invention may be for either a "prophylactic" or "therapeutic" purpose. When provided prophylactically, the agent(s) are provided in advance of any symptoms indicative of the organisms growth. The prophylactic administration of the agent(s) serves to prevent, attenuate, or decrease the rate of onset of any subsequent infection. When provided therapeutically, the agent(s) are provided at (or shortly after) the onset of an indication of infection. The therapeutic administration of the compound(s) serves to attenuate the pathological symptoms of the infection and to increase the rate of recovery.

The agents of the present invention are administered to the mammal in a pharmaceutically acceptable form and in a therapeutically effective concentration. A composition is said to be "pharmacologically acceptable" if its administration can be tolerated by a recipient patient. Such an agent is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient.

The agents of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby these materials, or their functional derivatives, are combined in admixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation, inclusive of other human proteins, e.g., human serum albumin, are described, for example, in *Remington's Pharmaceutical Sciences* (16th ed., Osol, A., Ed., Mack, Easton PA (1980)). In order to form a pharmaceutically acceptable composition suitable for effective administration, such compositions

**THIS PAGE BLANK (USPTO)**

mediating such effects are disclosed in *Remington's Pharmaceutical Sciences* (1980).

For example, a change in the immunological character of the functional derivativ , such as affinity for a given antibody, is measured by a competitive immunoassay. Changes in immunomodulation activity are measured by the appropriate assay. Modifications of such protein properties as redox or thermal stability, biological half-life, hydrophobicity, susceptibility to proteolytic degradation or the tendency to aggregate with carriers or into multimers are assayed by methods well known to the ordinarily skilled artisan.

The therapeutic effects of the agents of the present invention may be obtained by providing the agent to a patient by any suitable means (i.e., inhalation, intravenously, intramuscularly, subcutaneously, enterally, or parenterally). It is preferred to administer the agent of the present invention so as to achieve an effective concentration within the blood or tissue in which the growth of the organism is to be controlled.

To achieve an effective blood concentration, the preferred method is to administer the agent by injection. The administration may be by continuous infusion, or by single or multiple injections.

In providing a patient with one of the agents of the present invention, the dosage of the administered agent will vary depending upon such factors as the patient's age, weight, height, sex, general medical condition, previous medical history, etc. In general, it is desirable to provide the recipient with a dosage of agent which is in the range of from about 1 pg/kg to 10 mg/kg (body weight of patient), although a lower or higher dosage may be administered. The therapeutically effective dose can be lowered by using combinations of the agents of the present invention or another agent.

As used herein, two or more compounds or agents are said to be administered "in combination" with each other when either (1) the physiological effects of each compound, or (2) the serum concentrations of each compound can be measured at the same time. The composition of the

**THIS PAGE BLANK (USPTO)**



biological activity of the protein, while other agents may bind to a component of the outer surface of the organism blocking attachment or rendering the organism more prone to act the bodies nature immune system. Alternatively, the agent may be comprise a protein encoded by one of the ORFs of the present invention and serve as a vaccine. The development and use of a vaccine based on outer membrane components, such as the LPS, are well known in the art.

As used herein, a "related organism" is a broad term which refers to any organism whose growth can be modulated by one of the pharmaceutical agents of the present invention. In general, such an organism will contain a homolog of the protein which is the target of the pharmaceutical agent or the protein used as a vaccine. As such, related organism do not need to be bacterial but may be fungal or viral pathogens.

The pharmaceutical agents and compositions of the present invention may be administered in a convenient manner such as by the oral, topical, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes. The pharmaceutical compositions are administered in an amount which is effective for treating and/or prophylaxis of the specific indication. In general, they are administered in an amount of at least about 10  $\mu\text{g/kg}$  body weight and in most cases they will be administered in an amount not in excess of about 8 mg/Kg body weight per day. In most cases, the dosage is from about 10  $\mu\text{g/kg}$  to about 1 mg/kg body weight daily, taking into account the routes of administration, symptoms, etc.

The agents of the present invention can be used in native form or can be modified to form a chemical derivative. As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of

**THIS PAGE BLANK (USPTO)**

hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent, in the control of bacterial infection by modulating the activity of the protein encoded by the ORF. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition for use in controlling *Haemophilus* growth and infection.

### 5. Vaccine and Pharmaceutical Composition

The present invention further provides pharmaceutical agents which can be used to modulate the growth of *Haemophilus influenzae*, or another related organism, *in vivo* or *in vitro*. As used herein, a "pharmaceutical agent" is defined as a composition of matter which can be formulated using known techniques to provide a pharmaceutical composition. As used herein, the "pharmaceutical agents of the present invention" refers to the pharmaceutical agents which are derived from the proteins encoded by the ORFs of the present invention or are agents which are identified using the herein described assays.

As used herein, a pharmaceutical agent is said to "modulate the growth of *Haemophilus sp.*, or a related organism, *in vivo* or *in vitro*," when the agent reduces the rate of growth, rate of division, or viability of the organism in question. The pharmaceutical agents of the present invention can modulate the growth of an organism in many fashions, although an understanding of the underlying mechanism of action is not needed to practice the use of the pharmaceutical agents of the present invention. Some agents will modulate the growth by binding to an important protein thus blocking the

THIS PAGE BLANK (USPTO)

antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or DF.

5       Types of detection reagents include labelled nucleic acid probes, labelled secondary antibodies, or in the alternative, if the primary antibody is labelled, the enzymatic, or antibody binding reagents which are capable of reacting with the labelled antibody. One skilled in the art will readily recognize that the disclosed DFs and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well  
10       known in the art.

#### 4.       *Screening Assay for Binding Agents*

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents which bind to a protein encoded by one of the ORFs of the present invention or to  
15       one of the fragments and the *Haemophilus* genome herein described.

In detail, said method comprises the steps of:

- (a)       contacting an agent with an isolated protein encoded by one of the ORFs of the present invention, or an isolated fragment of the *Haemophilus* genome; and
- 20       (b)       determining whether the agent binds to said protein or said fragment.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or  
25       designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention.

THIS PAGE BLANK (USPTO)

5 *Introduction to Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. *et al.*, *Techniques in Immunocytochemistry*, Academic Press, Orlando, FL Vol. 1 (1982), V 1. 2 (1983), V 1. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

10 The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

15 In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

20 Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the DFs or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound DF or antibody.

25 In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a  
30 container which will accept the test sample, a container which contains the

**THIS PAGE BLANK (USPTO)**



The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues in which a fragment of the *Haemophilus influenzae* Rd genome is expressed.

5 The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. *et al.*, "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); 10 Jacoby, W.D. *et al.*, *Meth. Enzym.* 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as for immunoaffinity purification of the proteins of the present invention.

### 15 3. Diagnostic Assays and Kits

The present invention further provides methods to identify the expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using one of the DFs or antibodies of the present invention.

20 In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the DFs of the present invention and assaying for binding of the DFs or antibodies to components within the test sample.

25 Conditions for incubating a DF or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the DF or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the DFs or antibodies of the present invention. Examples of such assays can be found in Chard, T., *An*

THIS PAGE BLANK (USPTO)

protein (such as globulin or  $\beta$ -galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells.

Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al.*, *Exp. Cell Res.* 175:109-124 (1988)).

Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, A.M., *Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1984)).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to proteins of the present invention.

For polyclonal antibodies, antibody containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures.

The present invention further provides the above-described antibodies in detectably labelled form. Antibodies can be detectably labelled through the use of radioisotopes, affinity labels (such as biotin, avidin, etc.), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, etc.) fluorescent labels (such as FITC or rhodamine, etc.), paramagnetic atoms, etc. Procedures for accomplishing such labelling are well-known in the art, for example see (Sternberger, L.A. *et al.*, *J. Histochem. Cytochem.* 18:315 (1970); Bayer, E.A. *et al.*, *Meth. Enzym.* 62:308 (1979); Engval, E. *et al.*, *Immunol.* 109:129 (1972); Goding, J.W. *J. Immunol. Meth.* 13:215 (1976)).

**THIS PAGE BLANK (USPTO)**

As described here, the proteins of the present invention, as well as homologs thereof, can be used in a variety of procedures and methods known in the art which are currently applied to other proteins. The proteins of the present invention can further be used to generate an antibody which selectively binds the protein. Such antibodies can be either monoclonal or polyclonal antibodies, as well as fragments of these antibodies, and humanized forms.

The invention further provides antibodies which selectively bind to one of the proteins of the present invention and hybridomas which produce these antibodies. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing polyclonal and monoclonal antibodies as well as hybridomas capable of producing the desired antibody are well known in the art (Campbell, A.M., *Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth *et al.*, *J. Immunol. Methods* 35:1-21 (1980); Kohler and Milstein, *Nature* 256:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* 4:72 (1983); Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985), pp. 77-96).

Any animal (mouse, rabbit, etc.) which is known to produce antibodies can be immunized with the pseudogene polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of the protein encoded by the ORF of the present invention used for immunization will vary based on the animal which is immunized, the antigenicity of the peptide and the site of injection.

The protein which is used as an immunogen may be modified or administered in an adjuvant in order to increase the protein's antigenicity. Methods of increasing the antigenicity of a protein are well known in the art and include, but are not limited to coupling the antigen with a heterologous

**THIS PAGE BLANK (USPTO)**

TGTTTGGTAA GATTTTATGA AATCTTGAAC AGCTTTGCTG TCTTTGTTAT CGGTACGAGA 652980  
AACGATAATG TTCACATATG GAGAACTCTT ATCTTCTACA AATACACCGT CATCTTGAGC 653040  
ATTTAAGCCT ACTTGACCCG CATAAGTATT GTTTACTACA GCTAAATCAA CGTCGTCTAA 653100  
TGCGCGTGCC GCAACAGAAG TATCTACTTC AGTGATGTTT AATTTTTTCG GATTTTCAAC 653160  
AATATCTAAT ACAGTTGAAA GAAGGTTATT TGCATCTTTT AATTGATTA AACCTTGTTT 653220  
TTCAAGAAGA ATTAATGCAC GACCACGGTT TGTGGATCG TTAGGAACAA CAACTTTAGC 653280  
ACCGTCTTGT AATTCATTCA CATTTTIGAT TTTTITAGAA TAACCCGCTA ATGGATAGAC 653340  
GAAAGTATTA CCCACGATAA CTAAGTTATT TAAATTTTTC GCTTTTGCAT CTTCATCTAA 653400  
ATAAGGTTTA TGTTGCATTG CGTTTGCATC TAAATCACCT TTAGATACAG CTTCAATTTGG 653460  
TAATGCGTAG TCATTGAATT CAACGAATTG AACGTCTAAA CCATATTTTT CTTTAGCGAC 653520  
TTTTGCTGCA ATTTCTGCAA CTTGATGCTC TGGGCCAGAC ATCAGCCTA CTTTGATTTT 653580  
AAGCGGTGCT GCTGCAGCTT CAGGTTTTTT GTCTTCTTTA CAGCCTGTTA AAACGAGAGC 653640  
TGATGCGATT GCAGTGATTG CAAAAGTTG TTTTAATTTT ATAGGATTTT CTTCTGTTT 653700  
AAGTTAAGAG TTGCGTTTAA TTAACGATGA TCCACTTTTT TAGCCAGTGT ATCGCCGAGT 653760  
TTTTGGCTAA TCATAACGAA TAGCACAATA ATAATGGTTG CCACCCAAGT TACATAAGGC 653820  
ATATTGCGAT ATACGCCCGT AGTTGATAGC GAGGCTACCT AAACCACCGC CCCCTTGTGT 653880  
TCCTGCCATT GCAGAATAAC CAACTAAAGT GACTAGCGTA AGAGTAACGC CATTAATTAG 653940  
CGTAGGTAGA GCTTCTGACA AATAGAATTT ACGAACAATT TGCCATTTAG TAGCCCCCAT 654000  
TGCTTGAGCT GCTTCGGTTA AACCATTTGG AATTTCATT AGTGCAATTAG CAGTTAAGCG 654060  
AGCCACGAAT GGCATTGCAC AAATACTCAA TGGAATAATT GCTGCTGTTG TACCTAATAC 654120  
AGTTCCCACG ATGAAACGAG TTACAGGTAA TAAGATTAGG AGCAAATAA TAAATGGAAT 654180  
GGAACGCCCA ATATTAATAA TCGTGTTTAA CACAAAATGA GTGCGGTTAT TTTGTAAAAT 654240  
CTCATTTTTT CCAGTTAAAA AAGTCCATAT GCCAACAGGC ACGCCGACTA GTACAGCAAG 654300  
TAGGGTAGAT GCAAACTGA TATAAACAGT TTCATAAGTt GCGSTTGCGA CAACGCCCCA 654360  
CATTTGAGGT GTTAATTGCT GGCTGAATGT TGCCAAAAAA TCATTGAACA TAACCAAGCA 654420  
CCTCCACGCG CACGTTGTTT TCCATTAAAT AAACTTTGGT TTGCGTAATG GCATCTTCAT 654480  
CGCCTTCTAC TTCAGCAATG GTATAACCAA ATTTACAGCC TCCAGCATAA TCAATTTGTG 654540  
AGGTAAAGAT ACTGAGCTCC ACGCCAAAT TTTTGTATGC TTGAGAAAGT AACGGTGCAT 654600

THIS PAGE BLANK (USPTO)



GTGCTTATTT ATCTTGCTCA ACAGCTGGGT TTTGAAGACG ATATTATGTA TTTGATGCAT 651300  
 TACCCCGCTT ACGAAGAGCA AGATAGTGAA GTGTGGCGTT ATATCTCACA TACATTGGTA 651360  
 CATTATCTA ATTTGCATAT TTTGTTAAT CTGTCTTGGT TTTTATTTT CGGTGGAATG 651420  
 ATTGAGCGCA CTTTGGGCTC GGTAAATTA TTGATGTTGT ATGTTGTAGC ATCTGCCATA 651480  
 ACAGGATATG TGCAAAATTA TGTGTCTGGC CCCGCTTTCT TTGGGCTTTC TGGTGTGTA 651540  
 TATGCGGTTT TAGGTTATGT ATTTATACGT GATAAATTGA ATCATCATTT ATTTGATCTG 651600  
 CCAGAAGGTT TTTTACAAT GTTATTGGTG GGGATCGCAT TAGGTTTAT TAGCCCTTA 651660  
 TTTGGTGTG AAATGGGAAA TGCTGCACAT ATTTCAGGCT TGATTGTCCG CTTGATTGG 651720  
 GGATTTATTG ATAGCAAATT GCGCAAAAT TCGCTAGAGT AAGTCCGTTT AATTTGGATA 651780  
 AAGGATTCGC TATGAAACAA TCGTTACGCC ATCAAAAAT TATTAACTG GTGGAACAAT 651840  
 CTGGCTATTT AAGCACGGAA GAATTGGTTG CTGCCTTAGA TGTAGCCCT CAAACGATTC 651900  
 GTCGTGATTT GAATATTTTG GCGGAGTTAG ATTTAATTCTG CCGCCATCAC GGTGGTGCGG 651960  
 CATCGCCATC TTCTGCAGAA AATTCTGATT ACGTGGATCG TAAACAATTC TTTTCATTAC 652020  
 AAAAAAATAA TATTGCACAA GAAGTTGCGA AGTTGATTCC TAACGGTGCA TCGTTGTTTA 652080  
 TTGATATTGG TACGACGCCG GAGsTGTCG CCAATGCGTT GCTTGGTCAT GAAAACTCA 652140  
 GAATTGTGAC GAACAATCTG AATGCCGCTC ATCTTTTACG CCAAATGAA AGTTTTGATA 652200  
 TTGTCATGGC GGGCGGATCA TTACGAATGG ATGGTGGAAT TATCGGCGAA GCTACGGTAA 652260  
 ATTTTATTTT TCAATTTCTG CTAGATTTCTG GTATTTTAGG GATTAGTGCG ATTGATGCAG 652320  
 ATGGTTCATT ATTGGATTAT GATTACCATG AAGTACAAGT AAAACGAGCG ATTATTGAAA 652380  
 GTTCACGTCA GACCTTATTA GTGGCCGATC ACTCTAAATT TACTCGCCAA GCGATTGTTC 652440  
 GCTTGGGCGA ATTAAGTGAT GTGGAATATT TGTTTACAGG TGATGTTTCT GAGGGCATTG 652500  
 TCAATTATTT GAAAGAGCAG AAAACGAAAT TGGTTTTATG TAATGGTAAA GTGCGGTAAA 652560  
 AATATCTTTG ATTTTCTTAT TAATTTGTCTG ATAAAAGAAG CTATTACTCT TAATTTTGA 652620  
 TTTAATTGAA ATTTATCTAA AAGTAATCTA AATATTCTCT CGCTGAATTA TTGTCTTTTC 652680  
 TATAATTAAA CATTTCCAAG CGTAAAAGTA TCTTGCTTTT TTAATCAATT AAAGAAATTT 652740  
 CATATTTCTG TATTATCTAA TTATTTTITAG CTACCAATAT TCTTTAAAAC ATTCCGAGCT 652800  
 TAGATAAGAA AATCCCTTGC TGGGCAAGGG ATTTAGTAAT GGGTTAAAGT GCGGTAGAAA 652860  
 TTACCAACCT TTTACAACGC CATCTTTAAA GTGTTTTTGA CTTCTTGGT AAACCTCTTC 652920

**THIS PAGE BLANK (USPTO)**

***What Is Claimed Is:***

1. Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID NO:1, a representative fragment thereof or a nucleotide sequence at least 99.9% identical to the nucleotide sequence depicted in SEQ ID NO:1.

2. Computer readable medium having recorded thereon any one of the fragments of SEQ ID NO:1 depicted in Table 1a or a degenerate variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

3. The computer readable medium of claim 1, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.

4. The computer readable medium of claim 3, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.

5. A computer-based system for identifying fragments of the *Haemophilus* genome of commercial importance comprising the following elements;

a) a data storage means comprising the nucleotide sequence of SEQ ID NO:1, a representative fragment thereof, or a nucleotide sequence at least 99.9% identical to the nucleotide sequence of SEQ ID NO:1;

b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and

c) retrieval means for obtaining said homologous sequence(s) of step (b).

6. A method for identifying commercially important nucleic acid fragments of the *Haemophilus* genome comprising the step of comparing a database comprising the nucleotide sequence depicted in SEQ ID NO:1, a representative fragment thereof, or a nucleotide sequence at least 99.9% identical to the nucleotide sequence of SEQ ID NO:1 with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected.

7. A method for identifying an expression modulating fragment of *Haemophilus* genome comprising the step of comparing a database comprising the nucleotide sequence depicted in SEQ ID NO:1, a representative fragment thereof, or a nucleotide sequence at least 99.9% identical to the nucleotide sequence of SEQ ID NO:1 with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence comprises sequences known to regulate gene expression.

8. An isolated protein-encoding nucleic acid fragment of the *Haemophilus influenzae* Rd genome, wherein said fragment consists of the nucleotide sequence of any one of the fragments of SEQ ID NO:1 depicted in Table 1a or a degenerate variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

9. A vector comprising any one of the fragments of the *Haemophilus influenzae* Rd genome depicted in Table 1a or a degenerate variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

10. An isolated fragment of the *Haemophilus influenzae* Rd genome, wherein said fragment modulates the expression of an operably linked

open reading frame, wherein said fragment consists of the nucleotide sequence from about 10 to 200 bases in length which is 5' to any one of the open reading frames depicted in Table 1a or a degenerate variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

5           11. A vector comprising any one of the fragments of the *Haemophilus influenzae* Rd genome of claim 8.

12. An organism which has been altered to contain any one of the fragments of the *Haemophilus* genome of claim 8.

10           13. An organism which has been altered to contain any one of the fragments of the *Haemophilus* genome of claim 10.

15           14. A method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching 5' to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the fragments of the *Haemophilus* genome depicted in Table 1a or a degenerate variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

20           15. An isolated nucleic acid molecule encoding a homolog of any one of the fragment of the *Haemophilus* genome depicted in Table 1a, excluding the fragments of SEQ ID NO:1 depicted in Table 1b wherein said nucleic acid molecule is produced by the steps of:

- a) screening a genomic DNA library using any one of the fragments of the *Haemophilus* genome depicted in Table 1a as a target sequence;
- 25           b) identifying members of said library which contain sequences which hybridize to said target sequence;

c) isolating the nucleic acid molecules from said members identified in step (b).

5 16. An isolated DNA molecule encoding a homolog of any one of the fragments of the *Haemophilus* genome depicted in Table 1a, excluding the fragments of SEQ ID NO:1 depicted in Table 1b wherein said nucleic acid molecule is produced by the steps of:

- a) isolating mRNA, DNA, or cDNA produced from an organism;
- 10 b) amplifying nucleic acid molecules whose nucleotide sequence is homologous to amplification primers derived from said fragment of said *Haemophilus* genome to prime said amplification;
- c) isolating said amplified sequences produced in step (b).

15 17. An isolated polypeptide encoded by any one of the fragments of the *Haemophilus influenzae* Rd genome depicted in Table 1a or by a degenerate variant of said fragment, excluding the fragments of SEQ ID NO:1 depicted in Table 1b.

18. An isolated polynucleotide molecule encoding any one of the polypeptides of claim 17.

20 19. An antibody which selectively binds to any one of the polypeptides of claim 17.

20. A method for producing a polypeptide in a host cell comprising the steps of:

- 25 a) incubating a host containing a heterologous nucleic acid molecule whose nucleotide sequence consists of any one of the fragments of the *Haemophilus influenzae* Rd genome depicted in Table 1a or a degenerate

variant thereof, excluding the fragments of SEQ ID NO:1 depicted in Table 1b under conditions where said heterologous nucleic acid molecule is expressed to produce said protein, and

b ) isolating said protein.

**THIS PAGE BLANK (USPTO)**